AMENDMENTS TO THE SPECIFICATION

Please delete the figure from the paragraph beginning on page 2, line 18, and further amend the paragraph as follows:

Moreover, load balancing reduces average system response time. Average response time in a system is the most widely used metric in measuring the efficiency of a load-balancing algorithm. In general, the average system response time in a system with or without load balancing mechanism looks like the ~~following:~~ graph depicted in FIG. 6. From the ~~above~~ graph, it is clear that load balancing can reduce the average system response time, most of the time, in a significant way.

Please amend the section entitled "Brief Description of the Drawings" as follows:

~~Fig~~ FIG. 1 depicts a logical partitioning of clients and servers in a communication network.

~~Fig~~ FIG. 2 depicts another logical partitioning of clients and servers where there are more servers than clients.

FIG. 3 depicts a flow diagram for a sender initiated algorithm that may be used to perform load balancing.

FIG. 4 depicts a flow diagram for a receiver initiated algorithm that may be used to perform load balancing.

FIG. 5 depicts a communication network with a client/server environment where servers are grouped with corresponding time period for loading change according to various aspects of the invention.

FIG. 6 is a graphical depiction of an average system response time in a system with and without a load balancing mechanism.

FIG. 7 is a graphical depiction of a determination of a load assignment time period according to an aspect of the present invention.

Please delete the figure from the paragraph beginning on page 26, line 25, and further amend the paragraph as follows:

The calculation of the time period T depends on the knowledge about the server service pattern and client request arrival pattern. The time period T1 through Tk may be calculated either analytically or by approximation. Any of the time periods Ti, where "i" is equal to 1 through "k", may be calculated based on L_i, the load level (number of requests) of servers in server group G_i, L_{i+1}, the load level of servers in G_{i+1}, and (ρ = (average request arrival rate)/(average server service rate) = λ/μ, which results in $Ti = -1/\mu * \log(\rho/(L_{i+1}-L_i\rho))$. A graphical representation of a determination of the time period is ~~as following:~~ depicted in FIG. 7. The calculation of time period T as shown may change without departing from the main scope of the invention.

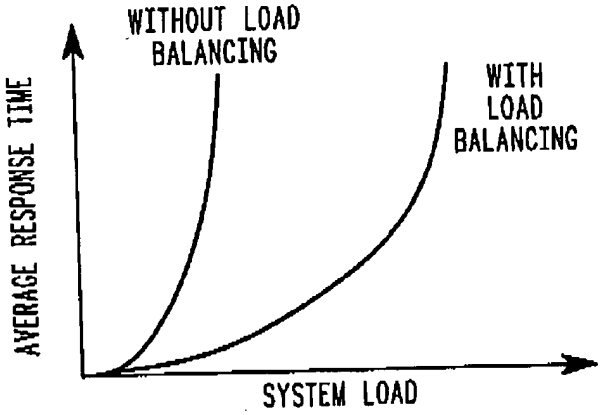

FIG. 6
—PRIOR ART—

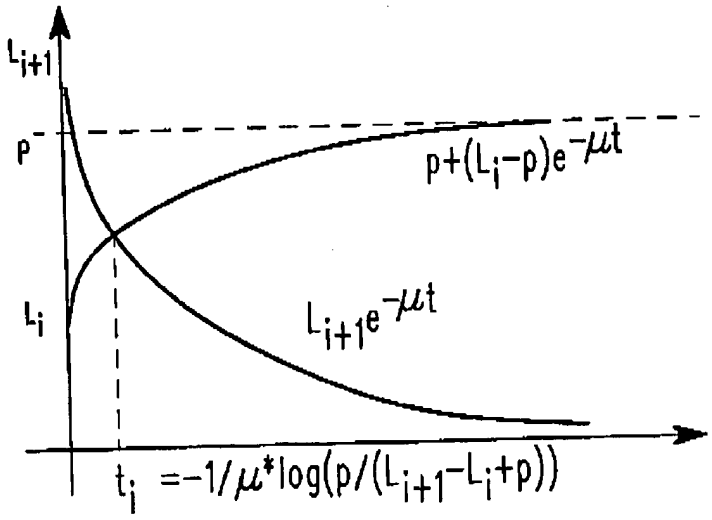

FIG. 7